Water Segmentation from SAR Images with the Presence of Speckle Noise

Roman Larionov [1] , Alexey Sennikov [1], Vladimir Khryashchev [1]

[1] P.G. Demidov Yaroslavl State University, Yaroslavl, Russia – r.larionov3@uniyar.ac.ru, alexeysennikov76@yandex.ru, v.khryashchev@uniyar.ac.ru

Keywords: Remote Sensing, Image Segmentation, Deep Learning, Speckle-Noise, Image Filtering, Despeckling

Abstract

Paper presents water segmentation algorithm in satellite SAR images. One of the purposes of water segmentation is flood monitoring and assessment its scale. Flood monitoring is complicated by the presence of severe weather conditions such as raining and cloudness. For this reason, C-band images were chosen which ignore atmospheric conditions and time of day. For the investigation, a dataset of 27 satellite images form Sentinel-1 with a spatial size of approximately 200 by 300 kilometers with a resolution of 10 meters per pixel was collected. U-ResNet-34, SegFormer_b5 and SegNeXt_l neural networks are used as segmentation algorithms. Training using balanced batch and augmentation invariance was used to improve the quality of the algorithms and the highest Dice value was equal 0.90. The paper also considers speckle noise filtering using a median filter and BM3D which allowed increasing the F1value by 0.01.

1. Introduction

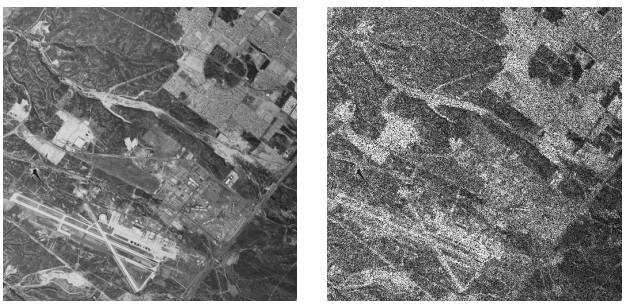
Remote sensing is the process of measuring the physical properties of remote objects by processing reflected or emitted energy. This task refers to the science of identifying features of the earth's surface and assessing the geobiophysical properties of objects using electromagnetic radiation as an interaction medium. Spectral, spatial, temporal and polarization indicators are the main characteristics that carry information about target objects (Moore, 1979).

Nowadays, the usage of remote sensing data has found a wide range of applications, for example, forest biomass assessment, agricultural and natural resource management, geology, and disaster prevention (Robertson, 2020). To solve such problems, segmentation algorithms are used that process satellite images using computer vision and machine learning methods.

Modern satellite constellations operate in a wide spectral range, including visible, infrared, short-wave infrared, and microwave. The quality of data in the first three ranges is highly dependent on atmospheric conditions, time of day, cloudiness, and other natural factors. Given these limitations and the low frequency of surveys, obtaining a satellite image of a local area at a given time may be unavailable. In this regard, the use of microwave data, in particular the C-band of wavelengths obtained by synthetic aperture radar (SAR) is in our field of interest. In this spectrum range, the properties of sensor signal penetration through clouds are evident, and the dependence of survey conditions on time of day and atmospheric conditions disappears (Erten et al, 2023).

Floods are among the most frequent and widespread natural disasters in the world, causing severe damage to life, property, infrastructure and the environment (Pai et al, 2019). With the development of technology, their monitoring and analysis using satellite data, computer vision and deep learning algorithms has recently become increasingly popular (Colin et al, 2022). The appearance of floods is accompanied by heavy rains and heavy clouds, which makes it impossible to recognize images in the optical range. To solve this problem, images with a synthetic aperture (SAR) are used, which allows monitoring in any weather and at any time of the day (Lv et al, 2022).

Due to some unique characteristics of SAR images, such as all-weather capability, cloud penetration, and specific object reflectivity, it has a significant advantage over other infrared or optical sensors. The quality of images obtained by SAR sensors is degraded by high levels of noise due to the its interference nature. The largest source of noise in a SAR image is speckle noise, which is caused by the interaction of the coherent radar beam and the relatively rough surface being imaged (Ko et al, 2022). Speckle noise makes typical image processing techniques applied to radar images very difficult. Speckle noise appears as a grainy or mottled pattern in images, making them difficult to interpret and analyse. An example of clear and speckled grayscale image of San Diego is shown in Figure 1.



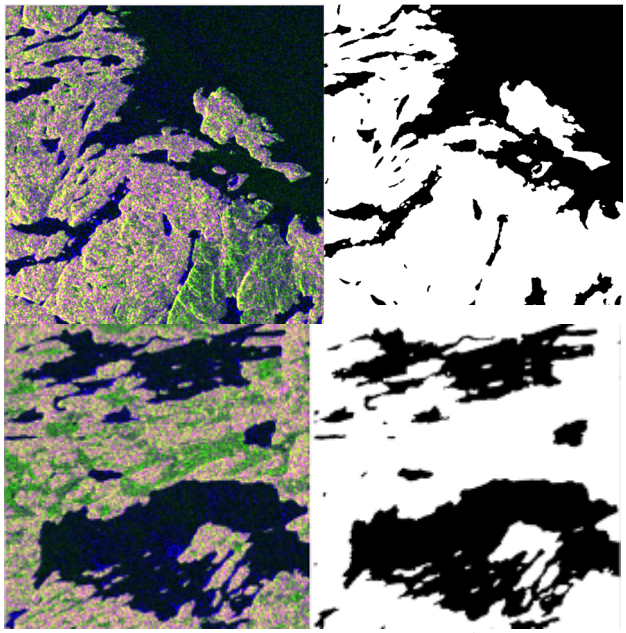
a b

Figure 1. Example of speckled image: original grascayle image (a), speckled grayscale images (b)

Effective despeckling SAR images is crucial for the task of satellite images segmentation. Many works are devoted to this topic. To eliminate speckle noise in SAR images, the use of filters such as the Lee filter (Lee, 1980), Kuan filter (Kuan et al, 1985) and Frost filter (Frost et al, 1982) is considered classic. These methods usually assume that the values of the image filtering result have a linear dependence on the original image, by searching for the corresponding combination of the central pixel intensity in a moving window with the average intensity in the filter window (Zhang et al, 2018). Thus, spatial linear filters provide a compromise between balancing in uniform areas and a constant full-pass identification filter in edge areas. The results confirmed that spatial domain filters successfully suppress speckle noise for some objects. However, due to the peculiarities of local processing, spatial linear filtering methods often do not allow to fully preserve edges and details, which leads to the following disadvantages (Yommy et al, 2015).

In addition to the spatial domain filters described above, wavelet theory has also been applied to reduce speckle noise (Starck et al, 2002). mainly used the ridgelet transform as a

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W5-2024-95-2024 | © Author(s) 2024. CC BY 4.0 License.



a b

Figure 2. Example of images 512x512 from train set: 2-band image (a), ground-truth mask (b)

component and realized the curved subbands using a discrete wavelet transform (DWT) filter bank to reduce the noise in the image. And for the case of speckle noise, researches in the work (Solbo et al, 2004) used the DWT of the logarithmically transformed image in homomorphic filtering, which is an empirical convergence in a self-adaptive strategy and is calculated in the Fourier space. The main disadvantages of this approach are preserving the mean backscatter in homogeneous regions, preserving details, and introducing artificial changes to the results such as ring effects (Lopez-Martinez et al, 2002).
In this work, we collected a dataset of satellite SAR images from the Sentinel-1 with water areas, rivers and floods, trained and tested neural networks for water segmentation, tested several filters for filtering speckle noise, and assessed filters on the quality of the segmentation algorithms.

2. Dataset

The collected dataset was downloaded from Copernicus (Europian Space Agency, 2014). It consists of twenty-seven 16-bit satellite images from the Sentinel-1 satellite of the Russian Federation taken in 2022/2024. Each image covered a spatial area of approximately 200 by 300 kilometers with a resolution of 10 meters per pixel. Satellite image included two bands – VV (vertical waves are transmitted and received) and VH (vertical waves are transmitted and horizontal waves are received). When forming images, preprocessing was carried out, consisting of removing thermal noise and noise along the edges of the image, radiometric calibration and geometric correction, and removal of speckle noise with a Lee filter (used only for training). Each image was assigned an expert binary mask identifying areas with water areas. 2 images were randomly selected for testing, all others were used for training. To form a training and test set, all images were divided into patches of size 512x512 pixels with a step of 256. Half-intersection is necessary to reduce the probability of the same part of the image water localization. The statistics of the resulting sets are presented in Table 1, examples of patches and corresponding binary masks are presented in Figure 2. For better figure visualization we added the third channel into images obtained by dividing VH by VV (VH/VV).

Number of images	Training set	Test set
Total	178817	13714
With water	97679	7202
Without water	72226	6512

Table 1. Dataset description.

3. Algorithms description

3.1 Neural networks architectures

In our research, we chose 3 neural network architectures: a fully convolutional neural network U-Net (Ronneberger et al, 2015) with ResNet-34 (He et al, 2016) as backbone, a convolutional network with attention SegNeXt_l (Guo et al, 2022) and MSCAN-L as backbone and a fully transformer network SegFormer-B5 (Xie et al, 2021) with a MiT-B5 as backbone.
The U-Net architecture (Ronneberger et al, 2015) follows the encoder–decoder structure. The aim of encoder is capturing the global context and feature extraction at different image scales. This is done only using convolutional and max-pooling operations. For this reason, U-Net belongs to the class of fully connected networks. The decoder's job is segmentation map formation based on the features extracted by encoder. It is done using two consecutive convolutional layers and upsampling operation. Moreover, to improve the network's ability of high-level features extraction, skip-connections are used to concatenate feature maps from encoder and the corresponding decoder blocks at the same scales. Skip-connections usage is the business card of U-Net and the main reason of its high-quality results in various tasks. Its architecture is shown in Figure 3.

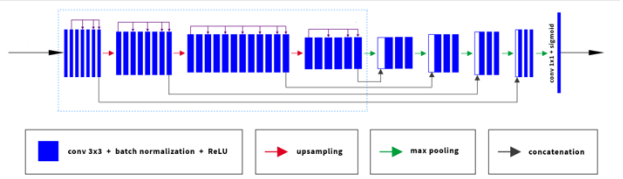


Figure 3. U-ResNet-34 neural network

SegNeXt (Guo et al, 2022) consist of four convolutional encoder and decoder blocks with feed-forward connections and residual base. Each encoder block named MSCA contains 5x5 convolution, three branches depth-wise convolutions for capturing global context and 1x1 convolution layer for building relationship between tensors of different shape. The decoder block consists of matrix decomposition operations for each level of extracted features. Illustration of SegNeXt is shown in Figure 4. SegNeXt is lightweight but still strong neural network.

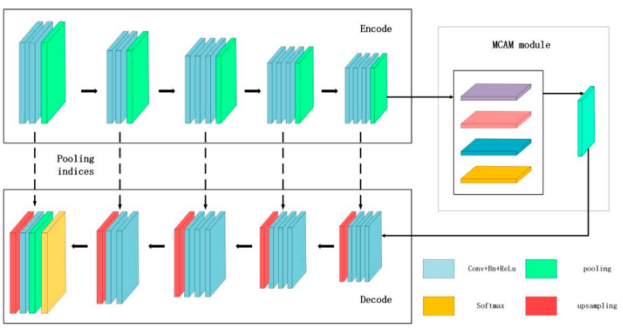


Figure 4. SegNeXt neural network

SegFormer (Xie et al, 2021) also belongs to encoder-decoder networks family. The encoder extracts global and local features from input image and consists of a patch embedding layer and 4 fully transformer blocks with self-attention. The patch embedding block splits the input image into 4x4 pixel patches and groups into fixed-size vectors, which are fed to the input of the transformer block. This block includes N-fold repeated self-attention layers and Mix-FFN layers that mix a 3x3 convolution and a feedforward network. The output of the transformer block is a fusion block that combines the features obtained in the previous step and reduces the dimensionality, obtaining a higher-level representation. The decoder is implemented on the basis of multilayer perceptrons and combines multilevel features produced by the encoder to form a segmentation mask. SegFormer framework is presented in Figure 5. SegFormer is usually called as U-Net from the transformer's world.

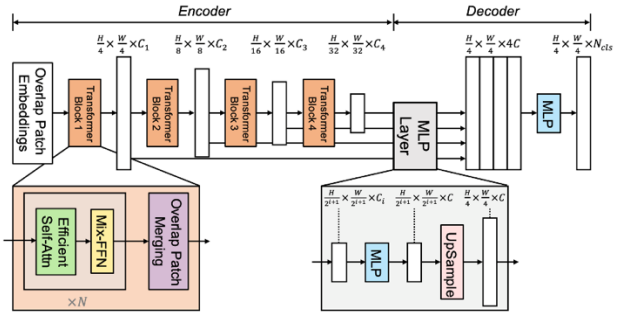


Figure 5. SegFormer neural network

3.2 Despeckling filters

In the current work, we focused on investigating classical despeckling filters without using deep learning methods: Lee filter, median filter and BM3D.

The Lee filter (Lee, 1980) is based on a statistical approach and uses local statistical properties of the image for adaptive smoothing. Its features are adaptability and edge preservation. The Lee filter is adaptive because it takes into account local statistical properties of the image, which allows it to effectively reduce speckle noise while preserving details and contrasts. Edge preservation is ensured by the fact that the Lee filter is based on local statistics, so it is able to preserve important edges and boundaries of objects that can be blurred by conventional smoothing methods.

The median filter (Huang et all, 1979) is a rank filter and one of the most popular filters for reducing noise in microscopic images. It belongs to the class of nonlinear filters, since its action is based on ranking the pixel values in the neighborhood of each pixel, rather than on a linear combination of these values . The median filter is often used in digital image processing, including in medicine, biology, and other fields where it is important to preserve image detail. The principle of the median filter is as follows. For each pixel of the image, a neighborhood is selected (usually square or rectangular in size, for example, 3x3 or 5x5 pixels). The intensity values of all pixels in this neighborhood are sorted in ascending order. After sorting, the median value is selected, which replaces the value of the central pixel of the neighborhood. The median value is the value that divides the set of sorted values in half, i.e. half of the values are less than the median, and the other half are greater. Its advantages are noise robustness and edge preservation. The median filter is most effective when filtering salt-and-pepper noise, which is somewhat similar to speckle noise, since single outliers in pixel intensity have virtually no effect on the median value.

The BM3D (Block-Matching and 3D Filtering) filter is a method for removing noise from images that uses block matching and 3D filtering. This method was proposed in 2007 (Dabov et al, 2007). BM3D is widely used due to its high efficiency in reducing noise while preserving image details. The BM3D filter works in 4 stages:

1. Grouping. At this stage, the image is divided into small overlapping blocks. For each block, similar blocks are found throughout the image using a similarity metric such as Euclidean distance. Similar blocks are combined into a three-dimensional array (cube).
2. 3D transformation and filtering. Three-dimensional transforms, such as the discrete cosine transform (DCT) or the wavelet transform, are applied to each three-dimensional array. The transformed coefficients are filtered, for example, using a hard or soft threshold, which allows noise to be removed.
3. Inverse transformation. After filtering, the coefficients are transformed back into a spatial representation. The processed blocks are returned to their original positions.
4. Aggregation. The processed blocks are combined to form the final image. Overlapping areas are averaged, which smooths out the transitions between blocks and reduces artifacts.

The advantages of BM3D are high efficiency, preservation of details and flexibility, which consists in the fact that the filter can be adapted to work with different types of noise. And its disadvantages are high computational complexity and parameters sensitivity (Yahya et al, 2020).

3.3 Implementation details

During the training process, we researched two training policies and its combination. The first policy follows from the task statement, which is binary segmentation. Random selection of images into a batch can lead to situations when the training batch contains only images with a water surface or only images without a water surface. This can lead to undesirable sharp jumps in the loss function. To eliminate this negative effect and stable training process, it is proposed to use a balanced batch: half of the images in the batch contain water and another half do not. We trained all developed networks with a random batch and a balanced one.

For the first policy we used the following types of augmentation:

- random brightness changes within 10%;
- mirror vertical and horizontal reflections;
- random rotations of 90, 180 and 270 degrees;
- random scale changes in scale within 10%;
- random rotations at angles up to 45 degrees;
- random cropping.

The second policy is usage of Augmentation Invariance (AI). The specificity of satellite image is its independence of specific rotation angles and presence of significant photometric distortions due to the satellite sensor tilt angle. Therefore, standard augmentation methods may have low efficiency in aerial image segmentation tasks. The approach to increase invariance to such transformations is that the original patch is augmented in two different ways, and the network predictions for both cases should match (Tavera et al, 2022). The illustrations of training iteration using Augmentation Invariance is shown in Figure 6. The first type of augmentation includes:

- mirror vertical and horizontal reflections;
- random scale changes in scale within 10%;
- random rotations at angles up to 45 degrees;
- random cropping.

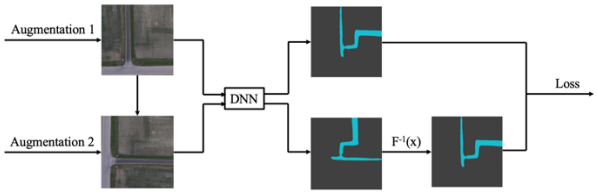


Figure 6. Illustration of Augmentation Invariance method

And the second type of augmentation consists of:

- random rotations of 90, 180 and 270 degrees;
- random brightness changes within 10%;

For training based on this approach, a common loss function was used that took into account the segmentation errors of the original image, the segmentation error of the transformed image, and the difference between the prediction from original image and prediction from the augmented image. Thus, the loss function can be expressed as:

$$L_{tot} = L_{seg}(x, y) + L_{seg}(A(x), A(y)) + \lambda L_{AI}(x, A(x)), \quad (1)$$

$$L_{AI}(x, A(x)) = \frac{1}{I}\sum_{i \in I}\left[f_i(x) - A^{-1}\left(f_i(F(x))\right)\right]^2, \quad (2)$$

where L_{seg} is a binary cross-entropy segmentation loss, L_{AI} is pixel-wise mean squared loss between augmented and non augmented images, A is combination of the first type augmentations, A^{-1} is inversion of geometric augmentations (reverse rotations of 90, 180 and 270 degrees), x is original input image, y is ground-truth and f_i is model prediction.

To develop our proposed methods we leverage mmsegmentation framework (MMSegmentation Contributors, 2020), which is based on PyTorch. We train every neural network on eight NVIDIA Tesla V100 GPUs with 16 GB RAM with batch of 2 per GPU. Thus, the total batch was 16 images. Training was carried out over 500K. we used AdamW optimizer. The base learning rate was set to 10^{-3} for U-Net and $6 * 10^{-5}$ for SegNeXt and SegFormer, the weight decay to 0.01. The learning rate was reducing following the poly law. Binary cross-entropy + dice loss was used as loss function for training with first policy and the combined loss from equation 1 was used for Augmentation Invariance.

3.4 Metrics

To evaluate the quality of developed algorithms, Dice coefficient and F1 score were used (Khryashchev et al, 2020):

$$dice = \frac{2I}{S}, \quad (3)$$

$$I = \sum_{x \in X,\ y \in Y} xy, \quad (4)$$

$$S = \sum_{x \in X,\ y \in Y} (x + y), \quad (5)$$

$$F1 = 2\frac{precision \times recall}{precision + recall}, \quad (6)$$

$$precision = \frac{TP}{TP + FP}, \quad (7)$$

$$recall = \frac{TP}{TP + FN}, \quad (8)$$

where I is union of predicted values and ground-truth, S is intersection, x is ground-truth pixel value, y is predicted value, TP is a number of true positive objects (object is TP if predicted and ground-truth polygon are intersected by more than 0.5 of IoU), FP is a number of false positive objects and FN is a number of false negative objects.

4. Experimental results

At the first stage we trained U-ResNet-34, SegNeXt-L and SegFormer using the first training policy. Test results for random and balanced training batch are presented in Table 2 and Table 3.

Model	Metrics	
	F1	Dice
U-ResNet-34	0.45	0.73
SegNeXt-L	0.40	0.87
SegFormer-B5	0.38	0.87

Table 2. Test results of developed segmentation algorithms using random batch.

Model	Metrics	
	F1	Dice
U-ResNet-34	0.44	0.74
SegNeXt-L	0.40	0.89
SegFormer-B5	0.39	0.89

Table 3. Test results of developed segmentation algorithms using balanced batch.

The highest value of Dice metric was equal 0.89 for the SegFormer-B5 while the best value of F1 metric was shown by U-net and was equal 0.45. That's connected with the great transformer's capability to extract high-level global features, thus it may miss small objects. We can also observe that using a balanced batch allowed us to increase the Dice value by 0.01 for all three neural networks, and by 0.02 of F1 for SegNeXt. Illustration of SegNeXt training loss values per iteration is shown in Figure 7.

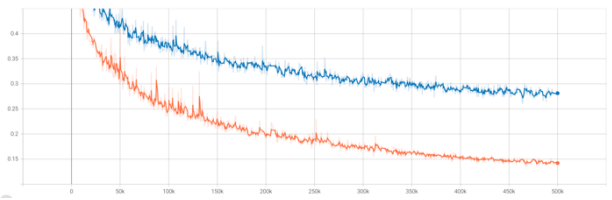


Figure 7. The training loss curve for SegNeXt: blue is random batch, red is balanced batch

Further, to increase the robustness of the algorithms, we trained every neural network using balanced batch and the augmentation invariance. It allowed us to increase the Dice and F1 values for all algorithms to 0.02. The increase in metrics is due to the fact that algorithms began to produce fewer false positives and also to better separate objects from each other. Results are shown in Table 4.

Model	Metrics	
	F1	Dice
U-ResNet-34	0.46	0.77
SegNeXt-L	0.42	0.90
SegFormer-B5	0.40	0.90

Table 4. Test results of developed segmentation algorithms using balanced batch and augmentation invariance.

The next step was study the effect of adding multiplicative speckle-noise to images with further filtering it on the quality of segmentation algorithms.
The noise model is described by the following expression:

$$noise\_img = img + img * gauss(mean, \mathrm{var}), \quad (9)$$

where img is an original image and gauss is normal distribution. For the investigation, the whole test set was noisy with a zero mean value and a variance value of 0.01; 0.05; 0.1; 0.5; 1. The test results of SegFormer-B5 are shown in Table 5. Speckled images with different variance values are presented in Figure 8.

Var value	Metrics	
	F1	Dice
0	0.39	0.86
0.01	0.39	0.90
0.05	0.39	0.90
0.1	0.39	0.90
0.5	0.39	0.90
1.0	0.37	0.90

Table 5. Segformer_b5 test results on noisy images.

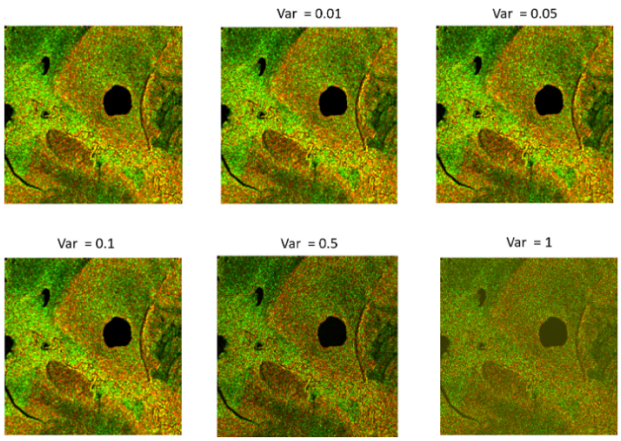
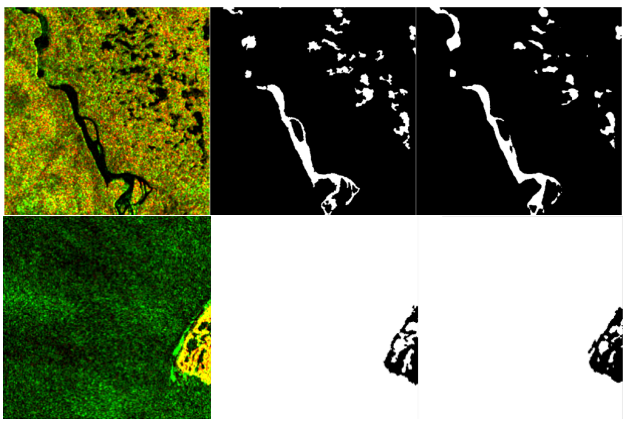


Figure 8. Images with different variance values if speckle noise

As follows from Table 5, with the noise increasing Dice value increases. That's caused with disappearing of false positives of few pixels size. Thus, the next testing was carried out only for big water objects whose size exceeds 1000 pixels. Then we investigated the effect of noise filtering on the quality of segmentation. We filtered the noisy test set using BM3D and median filters and calculated Dice and F1 metrics only for big objects. The results are presented in Table 6.
From Table 6 follows that SegFormer-B5 is resistant to speckle-noise, since the values of metrics with its addition and filtering change slightly. However, using the median filter allowed us to obtain a maximum value of 0.83 for the f1 metric. The segmentation results of developed algorithm with median filter post-processing are shown in Figure 9.

Filter type	Var value	Metrics	
		F1	Dice
No filtering	0	0.82	0.97
	0.05	0.82	0.97
	0.1	0.82	0.97
	0.5	0.81	0.97
	1.0	0.83	0.96
		F1	Dice
BM3D	0	0.82	0.97
	0.05	0.82	0.97
	0.1	0.82	0.97
	0.5	0.81	0.97
	1.0	0.83	0.96
		F1	Dice
Median	0	0.83	0.97
	0.05	0.83	0.97
	0.1	0.83	0.97
	0.5	0.83	0.97
	1.0	0.84	0.97

Table 6. Segformer_b5 test results on noisy images using BM3D and median filters.



a b c

Figure 9. Segmentation results of developed algorithm: images after median filter (a), ground-truth mask (b), segmentation result (c)

5. Conclusions

During the work, segmentation algorithms using deep learning methods were studied, as well as the problem of speckle noise, which is characteristic of SAR-images. The dataset of 27 satellite images from Sentinel-1 was collected and marked for research. We trained U-ResNet-34, SegNeXt-L and SegFormer-B5 for water segmentation using three policies and its combination: random batch, balanced batch, augmentation invariance. The quality of proposed methods was assessed using Dice and F1 metrics. The highest value of Dice was equal 0.90 using balanced batch and augmentation invariance. Adding speckle noise to a test set followed by digital filtering showed that the best despeckling filter was the median filter, which showed an increase in F1-score by 0.02.
Future research will look into developing despeckling filters based on deep learning algorithms.

References

Colin, A., Fablet, R., Tandeo, P., Husson, R., 2022: Semantic segmentation of metoceanic processes using SAR observations and deep learning. *Remote Sensing*, 14(4), 851. doi.org/10.3390/rs14040851

Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K., 2007: Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *EEE Transactions on Image Processing*, 16(8), 2080-2095. doi.org/10.1109/TIP.2007.901238

Erten, H., Bostanci, E., Acici, K., Guzel, M., Asuroglu, T., Aydin, A., 2023: Semantic Segmentation with High-Resolution Sentinel-1 SAR Data. *Applied Sciences*, 13(10), 6025. doi.org/10.3390/app13106025

Europian Space Agency, 2014: Copernicus Open Access Hub. https://scihub.copernicus.eu/ (10 July 2024)

Frost, V., Stiles, J., Shanmugan, K., Holtzman, J., 1982: A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(2), 157-166. doi.org/10.1109/TPAMI.1982.4767223

Guo, M.-H., Lu C.-Z., Hou, Q., Liu, Z., Cheng, M.-M., Hu. S.-M., 2022: SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. *Advances in Neural Information Processing Systems*, 35, 1140-1156.

He, K., Zhang, X., Ren, S., Sun, J., 2016: Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. doi.org/10.1109/CVPR.2016.90

Huang, T., Yang, G., Tang, G., 1979: A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(1), 13-18. doi.org/10.1109/TASSP.1979.1163188

Khryashchev, V., Larionov, R., 2020: Wildfire Segmentation on Satellite Images using Deep Learning. *2020 Moscow Workshop on Electronic and Networking Technologies (MWENT)*, 1-5. doi.org/ 10.1109/MWENT47943.2020.9067475

Ko, J., Lee, S., 2022: SAR Image Despeckling Using Continuous Attention Module. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3-19. doi.org/ 10.1109/JSTARS.2021.3132027

Kuan, D., Sawchuk, A., Strand, T., Chavel, P., 1985: Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2), 165-177. doi.org/10.1109/TPAMI.1985.4767641

Lee, J., 1980: Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2), 165-168. doi.org/10.1109/TPAMI.1980.4766994

Lopez-Martinez, C., Fabregas, X., 2002: Modeling and reduction of SAR interferometric phase noise in the wavelet domain. *IEEE Transactions on Geoscience and Remote Sensing*, 40(12), 2553-2566. doi.org/10.1109/TGRS.2002.806997

Lv, S., Meng, L., Edwing, D., Xue, S., Geng, X., Yan, X.-H., 2022: High-Performance Segmentation for Flood Mapping of HISEA-1 SAR Remote Sensing Images. *Remote Sens*., 14(21): 5504. doi.org/10.3390/rs14215504

MMSegmentation Contributors, 2020. MMSegmentation: Openmmlab semantic segmentation toolbox and benchmark. https://github.com/open-mmlab/mmsegmentation (10 July 2024)

Moore. G., 1979: What is a picture worth? A history of remote sensing. *Hydrological Sciences Bulletin*, 24(4), 477-485. doi.org/ 10.1080/02626667909491887

Pai, M. M. M., Mehrotra, V., Shreyas, A., Verma, U., Pai, R.M., 2019: Automatic segmentation of river and land in sar images: A deep learning approach. *2019 IEEE second international conference on artificial intelligence and knowledge engineering (AIKE)*, 15-20. doi.org/10.1109/AIKE.2019.00011

Robertson, L.D., 2020: Synthetic Aperture Radar (SAR) image processing for operational space-based agriculture mapping. *International Journal of Remote Sensing*, 41(18), 7112-7144. doi.org/10.1080/01431161.2020.1754494

Ronneberger O., Fischer P., Brox T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science. Springer*, 9351, 234–241. doi.org/10.1007/978-3-319-24574-4_28

Solbo, S., Eltoft, T., 2004: Homomorphic wavelet-based statistical despeckling of SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(4), 711-721. doi.org/10.1109/TGRS.2003.821885

Starck, J.-L., Candes, E., Donoho, D., 2002: The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6), 670-684. doi.org/10.1109/TIP.2002.1014998

Tavera, A., Arnaudo, E., Masone, C., Caputo, B., 2022: Augmentation Invariance and Adaptive Sampling in Semantic Segmentation of Agricultural Aerial Images. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1655-1664. doi.org/10.1109/CVPRW56347.2022.00172

Xie, E., Wang, W., Yu, Z., Anandkumar, A., 2021: Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34, 12077–12090.

Yahya, A.A., Tan, J., Su, B., 2020: BM3D image denoising algorithm based on an adaptive filtering. *Multimed Tools Appl* 79, 20391–20427. doi.org/10.1007/s11042-020-08815-8

Yommi, A., Liu R., Wu, A., 2015: SAR Image Despeckling Using Refined Lee Filter. *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, 260-265. doi.org/10.1109/IHMSC.2015.236

Zhang, Q., Yuan Q., Li, J., Yang, Z., Ma, X., 2018: Learning a Dilated Residual Network for SAR Image Despeckling. *Remote Sensing*, 10(2), 196. doi.org/10.3390/rs10020196